PRIMUS
Problems, Resources, and Issues in Mathematics Undergraduate Studies

ISSN: 1051-1970 (Print) 1935-4053 (Online) Journal homepage: http://www.tandfonline.com/loi/upri20

Roots of Linear Algebra: An Historical Exploration of Linear Systems

Christine Andrews-Larson

To cite this article: Christine Andrews-Larson (2015) Roots of Linear Algebra: An Historical Exploration of Linear Systems, PRIMUS, 25:6, 507-528, DOI: 10.1080/10511970.2015.1027975

To link to this article: http://dx.doi.org/10.1080/10511970.2015.1027975

Accepted author version posted online: 01 Apr 2015.
Published online: 01 Apr 2015.

PRIMUS, 25(6): 507–528, 2015

ISSN: 1051-1970 print / 1935-4053 online
DOI: 10.1080/10511970.2015.1027975

Roots of Linear Algebra: An Historical Exploration of Linear Systems

Christine Andrews-Larson

Abstract: There is a long-standing tradition in mathematics education to look to history to inform instruction. An historical analysis of the genesis of a mathematical idea offers insight into: (i) the contexts that give rise to a need for a mathematical construct; (ii) the ways in which available tools might shape the development of that mathematical idea; and (iii) ways in which students might make sense of an idea. In this paper, I discuss historic contexts that gave rise to considerations of linear systems of equations and their solutions, as well as implications for instruction and instructional design.

Keywords: Systems of linear equations, linear algebra, history.

1. INTRODUCTION

History provides a wealth of resources with the potential to inform the teaching and learning of mathematics [2,6,22]. Instructional insights can be gleaned from history by considering the contexts that gave rise to a need for a mathematical idea, the ways in which available tools might shape the development of that idea, and the ways in which students might make sense of that idea. Such insights can be particularly important for instruction and instructional design in inquiry-oriented approaches where students are expected to reinvent significant mathematical ideas. Sensitivity to the original contexts and notations that afforded the development of particular mathematical ideas is invaluable to the instructor or instructional designer who aims to facilitate students' reinvention of such ideas. This article explores the ways in which instruction and instructional design in linear algebra can be informed by looking to the historical roots of the subject.

Broadly speaking, this article is organized around a set of compelling examples from history that mark important conceptual developments in linear algebra's history (with an emphasis on developments relating to systems of

Address correspondence to Christine Andrews-Larson, Florida State University, 1114 West Call Street, Tallahassee, FL 32306, USA. E-mail: cjlarson@fsu.edu

linear equations and their solution sets), and that have the potential to inform instruction and instructional design. The discussion and analysis attends to the role of context, mathematical tools and representations, and the central ideas and driving questions that drove development. The work is structured around these issues because of their potential to inform the design and use of problem-solving tasks for students – particularly with regard to contextual framings for tasks, anticipating the role of tools and notation in affording students' productive engagement in tasks in ways that are likely to align with instructional goals, and identification of key mathematical ideas, particularly those that merit, demand and/or came about through reflective abstraction.

Three of the most surprising things I have learned from my excursion into the literature on the history of linear algebra have to do with Gaussian elimination. The first surprise was learning that the idea of systematic elimination that underlies Gaussian elimination preceded Gauss by over 2000 years – there is evidence that the Chinese were using an equivalent procedure to solve systems of linear equations as early as 200 BC [13,24]. The second surprise was that Gauss developed the method we now call Gaussian elimination without the use of matrix notation as it is commonly used in Western mathematics today. The third surprise was that Gauss developed the method we now call Gaussian elimination to find the best approximation to a solution to an overdetermined, inconsistent system of equations that had twice as many equations as unknowns.

This piece begins with a brief discussion of the theoretical underpinnings of this work, followed by a broad overview of several important historical developments in linear algebra. Next, the narrative highlights Gauss's work that gave rise to the method of solving systems of linear equations using what is now commonly referred to as Gaussian elimination; this is contrasted with the development of a remarkably similar procedure developed in ancient China $\sim$ 200 BC. The concluding discussion focuses on implications for instruction that emerge from this analysis, the contexts that give rise to a need for ideas relating to linear systems of equations, and the ways in which available tools and representations shaped the development of those ideas. This discussion informs the identification of central, underlying ideas and questions that drove the development of a coherent theory of systems of linear equations.

The need for a mathematical idea can prompt development when that need (e.g., through a problem or context) coincides with sufficient tools and notation to address the problem. Methods for solving systems of linear equations with unique solutions required less sophisticated notational tools and solution methods than systematically approximating solutions to inconsistent systems or solving systems with infinitely many solutions. Efforts to comprehensively characterize linear systems and their solutions grew into the theory of determinants; efforts to approximate solutions to inconsistent systems gave rise to Gaussian elimination. Significant advances in notation (namely the development of a convention for denoting a parameter) facilitated a historic shift from viewing solutions to linear systems as simply results of mathematical processes to viewing solutions to linear systems as mathematical objects in their own right. The contexts and notation that relate to these advances are detailed below, as are recommendations for helping students shift toward conceptualizing, as mathematical objects, the solutions to systems of linear equations.

2. THEORETICAL BACKGROUND

This work reflects the underlying view espoused by Freudenthal [8] that mathematics is an inherently human activity that takes place within and relative to social and cultural contexts. In this capacity, activity is said to be mathematical in nature when it aims to develop increasingly sophisticated and general ways of organizing, quantifying, characterizing, predicting, and modeling the world by either creating new mathematical tools for dealing pragmatically with challenging issues that exist within a social/cultural context, or by using the tools and language of the existing mathematical community to reason and problem solve [15].

I also draw on Sfard's [20] distinction between structural (object) and operational (process) conceptions of mathematical ideas. In the structural view, one conceptualizes a mathematical idea as an object that can be seen and manipulated as a whole, without detailing the process that gave rise to that object. Under the operational view, a mathematical object is thought of as "a potential rather than actual entity, which comes into existence upon request in a sequence of actions" [20, p. 4]. For example, when a child first learns about counting and numbers, he or she cannot conceive of the meaning of the number five without counting up to it. Thus, a child who must count up to the number five in order to be able to conceptualize it does not yet have a structural view of numbers, but rather exhibits only an operational understanding. Sfard offers compelling evidence on both the individual psychological level and the broad historical level that illustrates how mathematical objects (e.g., rational numbers, negative numbers, complex numbers) are often operational in their origin, positing that it is in fact the reification of a mathematical process that gives rise to a mathematical object. Five becomes a mathematical object in a child's mind when he or she comes to see it as a number that has meaning (e.g., it represents the cardinality of a set of five objects) that can be thought of separately from the underlying process of counting to five. Both process and object views are valued as important aspects of the development of mathematical ideas.

3. IMPORTANT DEVELOPMENTS IN THE HISTORY OF LINEAR ALGEBRA

Historians of mathematics differ on what they view to be the most important contributions to the history of linear algebra [6, 7, 14]. However, there seems to be consensus that the history of linear algebra is situated in two related points of view. One point of view is that the development of a coherent, comprehensive characterization of systems of equations and their solutions is seen as a driving, underlying force behind the subject. I refer to this as the "systems view." The other point of view is that, central to what we now consider to be linear algebra, is the development of a formal, axiomatic way of algebraically defining relations among and operations on vectors. I refer to this as the "vector spaces and transformations view." I consider both approaches to be central to linear algebra, but I find the distinction to be useful for contextualizing my analysis and discussion of the history of linear algebra. This paper focuses primarily on the "systems view" in that it focuses primarily on the development of characterizations of systems of linear equations and their solution sets rather than the development and study of vector spaces and their properties.

3.1. Determinants: System Solving Origins

With the exception of the solution methods developed around 200 BC in China, a limited amount of progress[1] in the development of a comprehensive theory of systems of linear equations and their solutions was made until the 1600s and 1700s when determinants emerged (separately) in both Japan and Europe [6,16]. Before the development of determinants, the ancient Chinese methods for solving systems of equations used counting boards to represent problem constraints in rectangular arrays and to specify the sequence of manipulations to be performed in order to solve a given system. This method is described in greater detail later in this paper.

In 1683, Japanese mathematician Seki Kowa developed a version of determinants as part of a method for solving a nonlinear system of equations [16]. This method was described in a way that relied on the positions of coefficients arranged in a rectangular array, an indicator of the strong influence of Chinese mathematics on Japanese mathematics. Mikami [16] recreated Seki's illustration from the original manuscript for 2 x 2 and 3 x 3 cases as shown below in Figure 1 with the following explanation:

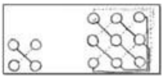

Figure 1. Mikami's recreation of Seki's diagrams.

[1]Methods for solving more than one equation through substitution and elimination were developed during the 1500s through the works of Cardano, Stifel, Buteo, Gosselin, and others. (See [12] for a treatment of this.)

> The dotted and real lines, or the red and black lines in the original manuscript, are used to indicate the signs which the product of the elements connected by these lines, will take in the development, the dotted lines corresponding to the positive sign and the real lines to the negative sign, if all the elements be positive. (p. 12)

A 2 x 2 example illustrates how this process yields a determinant as one would expect to see it today, although with a reversal of sign. Thus

becomes $-ad+bc$ because a and d are connected by a ''real'' (solid) line, so their product takes a negative sign whereas b and c are connected by a dotted line, so their product takes on a positive sign.

In 1750, Swiss mathematician Gabriel Cramer independently developed a way of specifying the solution to a system of linear equations as a set of closed-form expressions comprised purely of fixed but unspecified coefficients of the given system. (See [13] for a treatment of this.) Determinants can be seen in the denominators of these expressions, and Cramer generalized a method for their computation by leveraging the combinatorics of cleverly superscripted but unspecified coefficients. The representation Cramer [5] used to denote a system of equations with as many equations as unknowns is shown below in Figure 2. The upper case Z^is, Y^is, X^is, V^is denote coefficients while the lower case z, y, x, v denote unknowns. Note that Cramer's use of superscripts is consistent with contemporary use of subscripts in that they denote indexing (rather than exponentiation).

Cramer explicitly described closed-form solutions for such systems in 1, 2, and 3 unknowns; his solution for three equations and three unknowns is shown in Figure 3.

Although Cramer did not explain how this result was obtained, he offered a general rule for solving, framed in terms of the combinatorics of the superscripts. In this way, an $n \times n$ system is solved by forming n fractions, each

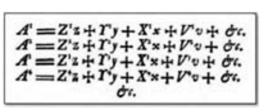

Figure 2. Cramer [5] denotes a system of equations.

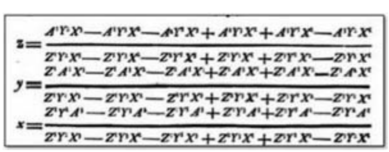

Figure 3. Cramer's solution to a system of three equations in three unknowns.

of which has *n!* terms in both the numerator and denominator. Katz [13] includes a translation of Cramer's description of how the denominators (i.e., determinants) are computed:

> Each term is composed of the coefficient letters, for example ZYX, always written in the same order, but the indexes are permuted in all possible ways. The sign is determined by the rule that if in any given permutation the number of times a larger number precedes a smaller number is even, then the sign is "+" otherwise it is "-". [13, p. 192]

Cramer noted that a zero denominator indicates that the system does not have a unique solution [13]. Furthermore, he specified that in the case that the denominator *and all* of the numerators are zero, the system will have infinitely many solutions, whereas when the denominator is zero but *any one* of the numerators is nonzero, the system will have no solution.

I argue that Cramer's approach would have been impossible without a shift in the use of algebraic notation introduced by French mathematician François Viète in the late-16th century. The use of literal symbols (e.g., using a variable such as x to represent an unknown fixed quantity, a fixed but unspecified quantity, or a varying quantity) was revolutionized in 1591 when Viete introduced the convention of using vowels to represent unknown quantities and consonants to represent quantities that are known but unspecified [3]. The particulars of his convention are no longer in use, but the distinction that arose was pivotal in shaping contemporary algebraic notation. This advance can be thought of as the specification of the idea of a parameter [3].

For both Seki and Cramer, the way in which their notational system structured the coefficients shaped the way the determinant was specified. Seki's articulation of the determinant leveraged the physical arrangement of the coefficients in order to specify the operations to be performed on those coefficients. Cramer's articulation of the determinant relied on the clever use of indexing in the coefficients to create a closed-form expression. It is plausible that Cramer's notational system affords an object view of systems of equations more strongly than does Seki's, per Cramer's observation about what the value of the determinant reveals about the solution to the system.

3.2. Euler's Inclusive Dependence

A second important development in the theory of systems of linear equations also took place in 1750. Swiss mathematician Leonhard Euler questioned whether a system of n linear equations with n unknowns has a unique solution, using the following system of equations as a counterexample: $3x - 2y = 5$ and $4y = 6x - 10$ [6]. This observation was made as part of a discussion of Cramer's paradox, which deals with the number of points of intersection of algebraic curves and the number of points needed to determine an algebraic curve. Euler gave additional examples with larger systems, and pointed out that it is possible for an equation to be "comprised of" or "contained in" others [6, p. 7]. Dorier tags this notion of Euler's with the term "*inclusive dependence,*" pointing out that our modern notion of linear dependence is more carefully defined and more sophisticated [6, p. 7]. The very language that an equation might be "contained in" or "comprised of" others suggests that early conceptions of linear dependence involved thinking of dependence as a property of an equation, or perhaps as a relationship between or among equations – rather than thinking of it as a property of a *set* of equations. Readers who have taught linear algebra will likely recall hearing students make analogous comments about certain vectors that are "dependent on" other vectors (rather than stating that those certain vectors are linear combinations of the other vectors, as the mathematician in us might hope). I posit that this is a natural, intuitive, informal way of conceptualizing notions of dependence, and that it is a useful conception that can serve as a basis for formalization. Historically speaking, Euler's observation certainly raised an issue that contributed to the development of our current conception of linear dependence.

Both Cramer's and Euler's work around systems of linear equations took place in the context of theory of curves, and although their observations differ in focus, both point to a central related issue. Cramer's comment identifies what the value of his denominators (i.e., the determinant) reveals about the uniqueness of the solution set to a square system of equations, whereas Euler's inclusive dependence points to the lack of a unique solution when there is redundancy of information in the equations themselves (perhaps implicitly assuming the same number of equations as unknowns). The relationship between Cramer's comment about determinants and Euler's observation about inclusive dependence is an important idea in any introductory linear algebra class; namely, that a consistent square system will have "inclusive dependence" (as Euler would describe it) if and only if it has infinitely many solutions (which is when Cramer's denominators or the determinant is zero).

Like Cramer's observation, Euler's observation is significant in that it marks a qualitative shift in perspective from his predecessors' process view of solutions – his reasoning identified properties of the system itself that held implications for the system's solution set. Euler's observation contrasts with earlier perspectives where mathematical reasoning focused primarily on the development and use of processes for solving linear systems, and shifts toward a notion of a solution to a system of equations as a mathematical object with its own properties (in this case, uniqueness).

3.3. Other Important Developments

In 1811, Gauss developed a method of least squares for finding the best approximate solution to an over-determined (and inconsistent) system of linear equations that had 12 equations and six unknowns. This system of equations used observational measurements to model the orbit of a celestial body. It was in this context that Gauss outlined the method of Gaussian elimination, which he developed without the use of matrices [10]. His discussion in this and earlier works reflects a complete understanding of the conditions under which a system has no solution, a unique solution, and infinitely many solutions. For instance, Gauss gave a detailed explanation of the relationship between elimination and the nature of the solution set to a system of linear equations in his 1809 *Theoria Motus* (*Theory on the Motion of Heavenly Bodies moving in Conic Sections*):

> We have, therefore, as many linear equations as there are unknown quantities to be determined, from which the values of the latter will be obtained by common elimination. Let us see now, whether this elimination is always possible, or whether the solution can become indeterminate, or even impossible. It is known, from the theory of elimination, that the second or third case will occur when one of the equations . . . being omitted, an equation can be formed from the rest, either identical with the omitted one or inconsistent with it, or, which amounts to the same thing, when it is possible to assign a linear function $\alpha\mathrm{P}+\beta\mathrm{Q}+\gamma\mathrm{R}+\delta\mathrm{S}+$ etc., which is identically either equal to zero, or, at least, free from all the unknown quantities. [9, p. 269].

Thus, we see that Gauss did understand that any system of linear equations can have no solution, a unique solution, or infinitely many solutions. Furthermore, he explained how one can identify the nature of the solution set based on the elimination process. He did not give a detailed explanation of Gaussian elimination in this 1809 work, but one appeared in an 1811 piece. Gaussian elimination, in the form in which it was originally proposed in 1811, is discussed in greater detail later in this paper. Recall that Gauss did not use matrices for notating systems of equations or performing elimination.

Unlike in Chinese and Japanese traditions, matrices did not come into use in Western mathematics until the late 1800s. The term matrix was coined in 1850 by the English mathematician James Joseph Sylvester, who was doing work with determinants.[2] In 1857, Sylvester's friend and colleague Arthur Cayley published his *Treatise on the Theory of Matrices* [4]. In this treatise, Cayley introduced matrices from a systems of equations point of view (as shown in Figure 4), and then proceeded to develop his theory of matrices as mathematical objects that could be added, multiplied, inverted, and so on.

It was not until the late 1800s that we see a shift from a systems view of linear algebra to a vector spaces view, and the work of Frobenius and Peano were

[2]Determinants were studied extensively in Western Europe prior to this, e.g., the description earlier in this paper regarding Cramer's work 100 years earlier. A more complete discussion of the history of determinants in Western Europe is beyond the scope of this paper but is summarized by Katz [13].

$\begin{aligned} X &= ax + by + cz, \\ Y &= a'x + b'y + c'z, \\ Z &= a''x + b''y + c''z, \end{aligned}$	$(X, Y, Z) = \left(\begin{array}{ccc} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{array} \right)(x, y, z)$

Figure 4. Cayley introduces matrices from a systems point of view [4].

notable in this shift. In 1875, Frobenius offered a definition for linear dependence and independence that worked for both equations and n-tuples, and that is equivalent to the modern standard definition. According to Dorier [6], this treatment of equations and n-tuples as equivalent objects in terms of linearity served as a significant step toward the contemporary treatment of vectors in linear algebra.

Kleiner discusses Italian mathematician Giuseppe Peano's 1888 formalization of the first modern definition of a vector space [14]. Katz argues that Peano's formal defnition of vector spaces 295 did not gain much attention or popularity until 1918 when they reappeared in Hermann Weyl's book *Space-Time-Matter* [13]. Here Weyl articulated an important relationship between a "systems view" and a "vector spaces and transformations view" of linear algebra:

> Weyl . . . brings the subject of linear algebra full circle, pointing out that by considering the coefficients of the unknowns in a system of linear equations in n unknowns as vectors, 'our axioms characterize the basis of our operations in the theory of linear equations'." [13, p. 204]

The remainder of this article focuses on the development of Gaussian elimination by Gauss in Europe in the 1800s, and the development of a remarkably similar procedure in ancient China. Both these accounts endeavor to describe the context(s) that created a need to solve systems of linear equations and the representations used to notate and manipulate those systems. The final section of the paper includes a discussion of the ways in which these historical insights might serve to inform instruction and instructional design.

4. EUROPEAN DEVELOPMENT OF GAUSSIAN ELIMINATION[3]

Gauss developed his method of Gaussian elimination in the context of astronomy. He was working to determine information about the elliptical orbit of

[3]In this section, I draw heavily on translations of Gauss's [9,10] original works as well as Kleiner's [14] and Althoen and Mclaughlin's [1] work in looking at the history of linear algebra.

$$
\begin{array}{ccccccccc}
0 & = & - & 183.93'' & + & 0.79363dL & + & 143.66dZ & + & 0.39493d\pi \\
 & & & & + & 0.95920d\phi & - & 0.18856d\Omega & + & 0.17387di; \\
0 & = & - & 6.81'' & - & 0.02658dL & + & 46.71dZ & + & 0.02658d\pi \\
 & & & & - & 0.20858d\phi & + & 0.15946d\Omega & + & 1.25782di;
\end{array}
$$

Figure 5. The first two of 12 linear equations in six unknowns [10].

an asteroid named Pallas, which was discovered in 1802 by Heinrich Olbers. At the time, Pallas was considered to be a planet. Gauss had a set of observational measurements, collected over a number of years, which could be used to determine the eccentricity and inclination of the orbit of Pallas. In order to do so, Gauss used his data set, together with then current theories of astronomy, to create a system of linear equations with six unknowns and 11 equations. (He actually began with 12 equations, but one of them seemed wildly inaccurate, so he discarded it.) The system carried conflicting constraints that arose due to measurement error. The first two equations from his system are shown in Figure 5.

In order to find a "best" approximation to a solution to this system of equations, Gauss developed a method of least squares. Gaussian elimination was developed as part of this method. Gauss explained the importance of considering the closest "solution" to systems that do not have a solution:

> If the astronomical observations and other quantities, on which the computation of orbits is based, were absolutely correct, the elements also, whether deduced from three or four observations, would be strictly accurate (so far indeed as the motion is supposed to take place exactly according to the laws of Kepler), and, therefore, if other observations were used, they might be confirmed, but not corrected. But since all our measurements and observations are nothing more than approximations to the truth, the same must be true of all calculations resting upon them, and the highest aim of all computations made concerning concrete phenomena must be approximate, as nearly as practicable, to the truth. But this can be accomplished in no other way than by a suitable combination of more observations than the number absolutely requisite for the determination of the unknown quantities. [9, p. 249]

In *Theoria Motus* [9] Gauss offered an overview of his least squares method, and a much more elaborate explanation of the Gaussian elimination portion of this method is given in his 1811 piece *Disquisitio de Elementis Ellipticis Palladis* [10]. Here, Gauss first defined functions (V^i, $i = 1, \ldots, \mu$, where μ is presumed to be a positive integer) of a finite-valued positive integer v number of unknowns ($p, q, r, s, \ldots$), and treated the observations as function values (M^i). Note that Gauss's use of a superscript is consistent with the contemporary convention for subscripts; his superscripts are used for indexing, rather than to indicate that a function or variable is being composed or exponentiated. When $\mu > v$, he noted that "an exact representation of all the observations would only be possible when they were all absolutely free from error . . . this cannot, in the nature of things, happen" [9, p. 254]. Gauss argued that the most probable values of the unknowns are those such that the sum of the squares of the differences between the computed and observed values of the functions (i.e., the sum of the squares of the errors e^i defined as $e^i = V^i - M^i$) is minimized. By expressing the functions in general linear form ($V^i = n^i + a^i p + b^i q + c^i r + d^i s + \ldots$ for real numbers $n^i, a^i, b^i, c^i, d^i \ldots$ with $i = 1, \ldots, \mu$) and noting that the sum of squares of the errors is minimized when the partial derivatives (with respect to unknowns $p, q, r, s, \ldots$) are all zero, Gauss obtained a system of equations in terms of the errors e^i, each of which can be expressed in terms of unknowns $p, q, r, s, \ldots$ Rewriting this system of equations in terms of p, q, r, s, Gauss subsequently described how the first variable p can be eliminated from the system of equations. He then described how one can continue eliminating one variable at a time until only one remains, at which point one could determine the value of the single unknown quantity and perform back substitution to determine the values of the other unknowns.

The central aspect of this process that is relevant to the contemporary treatment of Gaussian elimination is the sequential use of substitutions performed in such a way that one variable is removed with each step of the process until only one variable in one linear equation remains. The value of this single variable can then be determined from the equation, and the value is then substituted into the previous equation with two unknowns to solve; this process is repeated until all unknown values have been found [1].

5. SOLVING LINEAR SYSTEMS IN ANCIENT CHINA

The *Nine Chapters on the Mathematical Art* is an ancient Chinese text comprised of 246 problems and solution methods. It is believed to have been produced sometime between 200 BC and 50 AD. An earlier version of the text was burned during the reign of Emperor Ch'in Shih Huang, a controversially tyrannical ruler credited with the unification of China as well as construction of the Great Wall of China [24]. The problems in this text arose from contexts such as field measurement (which gives rise to the development of geometry, fractions, and square and cube roots); trade, commerce, and taxation (which give rise to the development of ratios, proportions, and systems of equations); and distance–rate–time problems.

5.1. China's Mathematical Toolbox ∼200 BC

In order to contextualize the mathematics that appears in the *Nine Chapters*, it is important to consider the mathematical tools and ideas that the Chinese had at their disposal at the time it was written. One of the most prominent mathematical tools in common use around 200 BC in China was the counting board, on which counting rods made of bamboo or ivory were arranged in rectangular arrays so that various calculations could be performed [21]. Counting rod arithmetic was the central method of calculation in Chinese mathematics beginning around 500 BC and continuing until it was gradually replaced by the abacus between 1368 and 1644 AD [21]. Common calculations included addition, subtraction, multiplication, and division. Standard algorithms for these calculations leveraged the structure of the base-10 system, much like common algorithms of today, although the procedures looked rather different than today's standard column arithmetic.

In addition to the use of counting boards and a base-10 number system, the Chinese also made use of positive and negative integers, as well as fractions. They did not use literal symbols to represent unknown or unspecified quantities, nor did they use a system of axiomatic deductive logic.

5.2. Linear Systems in the Nine Chapters

In analyzing the ancient Chinese methods for solving linear equations, Shen, Crossley, and Lun [21] noted that the Chinese had as many as seven different solution methods. There is some amount of overlap among these methods, and several of them only work in systems with one or two equations. In what follows I focus on only one of these methods: the more general method for solving systems of linear equations discussed in the problems in Chapter 8, whose title can be translated as "Rectangular Arrays."

Chapter 8 contains 18 problems, all of which correspond to linear systems with between two and six unknown quantities. With one exception, all of the problems have a unique solution with the number of constraints (equations) equaling the number of unknown quantities. The only exception to this was a problem whose solution set had one degree of freedom, which the author dealt with by adding a reasonable assumption to the context in the description of the solution.

5.3. Solving Linear Systems with Rectangular Arrays: An Example

A look at the translation of the *Nine Chapters* produced by Shen et al. [21] offers some insight into the types of contexts that gave rise to systems of linear equations in ancient China. Thirteen of the eighteen problems in Chapter 8 draw on agricultural contexts, dealing with quantities of livestock or grain (by number, weight, volume, or cost). The others contexts range from practical (pulling forces of horses, amounts of water used by families sharing a communal well, and amounts of chicken eaten by people based on their social class) to more riddle-like (combinations of different types of coins, weights of sparrows and swallows). These are evidence of the existence of a common currency for trade, the domestication of horses for use as beasts of burden, and the class structure of Chinese society around 200 BC. The first problem in Chapter 8 reads:

> Now given 3 bundles of top grade paddy, 2 bundles of medium grade paddy, [and] 1 bundle of low grade paddy. Yield: 39 dou of grain. 2 bundles of top grade paddy, 3 bundles of medium grade paddy, [and] 1 bundle of low grade paddy, yield 34 dou. 1 bundle of top grade paddy, 2 bundles of medium grade paddy, [and] 3 bundles of low grade paddy, yield 26 dou. Tell: how much paddy does one bundle of each grade yield. [21, p. 399]

In order to fully understand the context of this problem, it is important that the reader understand that paddy is grain and that dou is a unit used for measuring volume. If we were to rephrase the first sentence in a more contemporary way, it might read “A combination of 3 bundles of high-quality grain, 2 bundles of medium-quality grain, and 1 bundle of low-quality grain will yield 39 barrels of flour.”

The reader is instructed to begin laying down counting rods vertically in the far right column, in correspondence with the numeric values shown below in Figure 6(a) and manipulated as in Figure 6(b). The steps to follow are remarkably similar to what you might find in an undergraduate textbook describing the steps of Gaussian elimination, albeit with the conventional use of rows and columns switched.

Once the counting board is as shown in Figure 6(c), the Array (Fangcheng) Rule states, “. . . the low grade paddy in the left column is the divisor, the entry below is the dividend. The quotient is the yield of low grade paddy” [21, p. 399]. This gives that the yield (per bundle) of the low-grade paddy is 99/36 or 2 3/4 dou. The rule proceeds to describe how to use the yield per bundle of low-grade paddy with the middle column to obtain the yield per bundle of medium- and high-grade paddy, respectively. Thus, we see that the Chinese process was markedly similar to Gaussian elimination as it is now commonly taught, in that all entries above the diagonal of the coefficient matrix are eliminated, and then back substitution is performed to compute the values of the unknowns.

1	2	3
2	3	2
3	1	1
26	34	39

a. Initial arrangement of rods

1		3
2	5	2
3	1	1
26	24	39

b. “Multiply the middle column throughout by the superior grain on the right and continuously subtract it.” (van der Waerden, 1983, pp. 47–48)

		3
	5	2
36	1	1
99	24	39

c. Final arrangement of rods

Figure 6. Chinese method of solving using counting boards, depicted with Hindu-Arabic numerals.

5.4. Discussion

In the Chinese method, the quantities do not function as coefficients in a system of equations as we would conceive of them today. For instance, we would likely express the far right column as the equation $3x + 2y + z = 39$, where x is the yield of one bundle of top-grade paddy, y is the yield of one bundle of medium-grade paddy, and z is the yield of one bundle of low-grade paddy. I argue that coefficients of three, two, and one can be conceptualized in two ways. Based on the problem statement, they are obviously the given numbers of bundles of each quality of grain. However, these coefficients can also be viewed as the weights on the unknown yield rates x, y, and z. This latter conceptualization highlights the multiplicative relation between the number of bundles and the yield per bundle, as well as the additive relationship between the yields from each type of grain in each given combination. I contend it is likely more conceptually difficult for students (though not unimportant) to represent this situation with a system of equations than it would be for them to represent it with a counting board, as the counting board may mask the need to explicitly contend with and coordinate these additive and multiplicative relationships.

In using rectangular arrays to solve linear systems of equations, the central demand on the mathematical thinker is that he or she determines a solution to the system of equations. Mathematical activity is thus focused on the process of finding a solution. I argue that, in Sfard's [20] framing, this is an example of a context that promotes a process view of a solution to a system of linear equations. Furthermore, it seems that the way in which systems were represented afforded this intuitive, elimination-based problem-solving approach while also constraining the opportunity to view these linear systems or their solutions as mathematical objects in their own right.

6. PEDAGOGICAL RECOMMENDATIONS

History has the potential to inform instruction in a variety of ways. In particular, it can offer insights into difficulties students are likely to encounter, and it can serve as a source of inspiration for the development of tasks that create a need for the types of mathematical ideas we want students to learn. With these ideas in mind, in this section I discuss implications for instruction that emerged from my analysis. I begin with some comments relating to the issue of determinants discussed early in this paper. The remainder of my discussion is organized around the contexts that gave rise to a need for ideas relating to linear systems of equations, the ways in which available tools and representations shaped the development of those ideas, and the identification of central, underlying ideas and questions that drove the development of a coherent theory of systems of linear equations.

6.1. Instructional Implication: Reinventing Determinants

Historically, determinants were developed to help express solutions to systems of linear equations in terms of their coefficients [6]. Although it is not the case that the idea of determinants emerged from an explicit goal of computing a single value that would reveal whether a system had a unique solution, such a framing is potentially useful from a pedagogical point of view. Such a framing not only pushes toward an object view of systems of equations, it also pushes students to think about relationships between the coefficients in a system of equations and the nature of the system's solution set. Instructionally, one might present an opportunity for students to reinvent the notion of a determinant using the framing suggested in Figure 7. A similar approach to reinventing determinants has been used in the context of an inquiry-oriented differential equations course [19].

One way to see if the system in Figure 7 has a unique solution is to determine if the lines are not parallel, which is easily done by putting both equations in slope-intercept form (i.e., the form $y = mx + b$ commonly used in high school algebra). This is easily related to the requirement that $ad \neq bc$, or equivalently, that $ad\text{-}bc \neq 0$. In this way, determinants can be conceived of as *a tool for determining if a system of equations has a unique solution*.

It is worth noting that in three or more dimensions, the mathematics becomes more complicated because one must deal with linear combinations and not just scalar multiples (much like when dealing with span and linear independence). Katz [13] suggests Maclaurin's 1729 approach if students are to derive determinants for a 3×3 system: "given three equations in three unknowns *x, y, z,* he solved the first and second equations for *y* (treating *x* as a constant), then the first and third equations, then equated the two values and found *z*" (p. 193).

6.2. On Context

In ancient China, the contexts that gave rise to a need for linear systems came largely from agriculture and trade. Although the solutions to these systems were often fractional values, the problems tended to be constructed with integer-valued constraints. Gauss, on the other hand, worked in contexts that

Determine under what conditions the following system of equations has a unique solution for *x,y*. Assume that all quantities are real-valued.

$$\begin{cases} ax + by = k_1 \\ cx + dy = k_2 \end{cases}$$

Figure 7. Possible task for student reinvention of determinants.

required non-integer-valued coefficients, and he explicitly discussed the numbers of possible solutions to linear systems and the conditions under which each would occur. He also discussed in detail the need for increased accuracy provided by using multiple measurements that acknowledge error, and the way in which this created a system with conflicting constraints for which a "best" solution was needed.

Gauss's work points to the importance of seriously considering meaningful contexts in which the number of constraints (equations) exceeds the number of unknowns – and the importance of not simply dismissing such systems as having no solution. In the case where a system of equations with no solution arises from a meaningful context, there is likely a need to find the best approximation to a solution.

In addition to my recommendation, per the work of Gauss, that inconsistent systems receive their due attention in realistic contexts, I suggest that instructors not take it as immediate or obvious that the number of (independent) constraints needs to equal the number of variables in order to have a unique solution. It is not necessarily obvious to students why the number of (independent) equations must match the number of unknowns in order for a linear system to have a unique solution. For these reasons, I posit that instructors might precede the question "How do we solve linear systems of equations?" with "What does it mean to be a solution to a system of equations?" Subsequent classroom discussions about the following issues could offer insight into how students think about processes for solving system of equations.

1. How does the relationship between the number of unknowns and the number of constraints (equations) affect the solution to a system of linear equations?
2. How can we detect and account for conflicting constraints and/or hidden redundancies when counting the constraints described in the previous question?
3. How do we know, when we manipulate systems of equations or perform row/column operations, what information is changed, and what is left the same? (For instance, if we say two augmented matrices are row equivalent, what is equivalent about them, and how do we know that aspect of the system was unchanged by the row operations we performed?)

These questions point to a core set of ideas about systems of linear equations and their solution sets that are non-trivial for students, and are much richer in terms of theory than what a procedurally focused treatment of Gaussian elimination might entail. There is a rich history with a variety of contexts that have contributed to the development of the current theory of the nature of linear systems of equations and their solutions – and this set of ideas is a foundational part of a complete understanding of linear algebra.

6.3. On Tools and Representations

Mathematical tools impact the way in which ideas are notated, represented, and conceptualized. For instance, the use of counting boards in China facilitated a shift to the base-10 numeration system from an earlier system in which the location of digits did not indicate their magnitude, and counting boards clearly affected the way in which systems of equations were represented and manipulated so as to find solutions. Gauss's use of literal symbols, which distinguished unknown quantities from unspecified but known quantities, lent itself to the use of repeated substitutions – a crucial element Gauss's 1811 description of his method for solving linear systems. Differential calculus also served as an important mathematical tool in Gauss's development of his method of least squares, where his need for Gaussian elimination arose. Another example that supports this claim about the importance of mathematical tools is seen in Seki's ''matrix-like'' arrangement of terms to develop determinants as compared with Cramer's clever use of indexing for his more combinatorial description of determinants.

Pedagogically, this points to the importance of the selection and framing of mathematical tasks and questions, as well as the notation, representations, and tools used in the posing of those tasks. For instance, consider the Chinese use of rectangular arrays on counting boards to solve what would now be described as linear systems of equations. The representation is tightly tied to the quantities given in the real-world context, and the manipulations of the columns are easily and intuitively justified in a way that ties directly to the problem context. Coherence between problem context and representational tools can support students' problem-solving efforts. On the other hand, representations that are too tightly connected to specific contexts and needed manipulations may lend themselves more readily to process views as argued previously in this paper. Process views are developmentally important, but when there is a need to shift to an object view, a shift in notation may help facilitate this change in perspective.

7. FINAL REMARKS

So what is one to do with all this history? How might it inform the teaching of linear algebra? Some would argue that asking students to work directly with the contexts from history is a productive route; I would argue that this depends on whether this approach aligns with the learning goals of a particular course. The instructional design theory of Realistic Mathematics Education suggests that an effective model of supporting students' learning is by providing them with opportunities to work first in real-world contexts, and to then gradually shift away from those specific contexts, pushing students to conjecture what can be generalized and supporting them in shifting toward notation that captures those generalizations [11].

One might then, for instance, draw on the problems such as those given in the *Nine Chapters* as contexts for the teaching of row reduction (likely under the more modern convention in which rows correspond to equations). However, these problems tend to focus on systems with unique solutions. Although there are some mechanics to learn in row reduction, the insight gleaned from the systematic use of elimination and how that can extend to situations in which a linear system does not have a unique solution is an important conceptual learning goal, and one that seems to be significantly more troublesome for students than solving a system with a unique solution.

There might also be value in drawing on Gauss's context to teach about either the development of Gaussian elimination or ways of approximating solutions to inconsistent systems. However, Gauss's approach to the former is heavily embedded in a broader set of questions that are beyond the scope of a typical introductory linear algebra course. The latter, the idea of learning about ways of approximating solutions to inconsistent systems, is a topic with many relevant contemporary applications, but is often left as a special topic to be covered as time allows. This is likely due to the fact that, although highly relevant to a number of applied contexts, such approximation methods are not crucial for developing subsequent ideas in an introductory linear algebra course.

Drawing on my experiences and those of my colleagues interested in issues surrounding the teaching and learning of linear algebra, a pervasive source of student difficulty is describing and making sense of the solution sets to systems of linear equations which have infinitely many solutions. These difficulties are likely to impede students' learning of other important ideas in linear algebra (e.g., describing non-trivial null spaces, finding and interpreting eigenvectors). Anecdotally, students are often able to describe a solution as something that, when "plugged in" for the variables, creates a true statement or set of statements; however, when the solution is not a single value for each variable, interpretation becomes problematic. Student questions that have arisen in class discussions around solutions to systems of equations with infinitely many solutions (after solving systems using substitution and elimination methods, but prior to instruction on row reduction) include: "How do you know if three planes intersect in a plane or a line?''and ‘‘ How do you know how many parameters there are (for describing the solution set)?'' and "How do you know which variables can be parameters (for describing the solution set)?" The latter two questions highlight the value of a systematic approach to solving systems of linear equations that standardizes the choice of parameters (e.g., through row reduction).

Viewed through the lenses of history and mathematics education research, student difficulties with making sense of solution sets that are described using parameters are perhaps unsurprising. Historically, mathematical understanding of linear systems with unique solutions predated our understanding of those with infinitely many solutions. Indeed, notational tools were needed to accurately and concisely characterize solutions in the latter case. However, the literature on student thinking would suggest that it is precisely these notational tools that are one significant source of student struggle. It is well-documented that the varied use and interpretation of literal symbols (e.g., as unknowns, as variables, as parameters) is a source of difficulty for students at the secondary and tertiary levels [17, 23]. The fact that literal symbols are needed to describe solution sets in the case of systems with infinitely many solutions suggests that making sense of such solution sets would be a source of challenge for students. An example, such as that given by Euler for illustrating inclusive dependence, could prove pedagogically useful for helping students make sense of linear systems with infinitely many solutions as well as the conventions for describing their solution sets (both geometrically and algebraically).

Mathematical ideas do not develop until there is an intellectual need for them, as well as a sufficient set of notational tools to reason effectively in the context of that need. A need for a solution to a system of linear equations is rather natural, but the idea that some systems do not have unique solutions is perhaps less natural. As such, I argue that in addition to working in contexts with unique solutions, students need opportunities to work with contexts where an approximate solution to an inconsistent solution is needed, as well as contexts where they must make sense of parameters. Inconsistent systems could be taken from examples, such as that of Gauss, or from any other number of applied examples that are commonly used for systems of equations problems in high school and college textbooks (e.g., problems where one must relate quantities of particular kinds of food to number of calories from carbohydrates, fats, and proteins). Possani et al. [18] detail how a mathematical modeling problem set in a traffic flow context can be used support students' learning about parameterizations of linear systems with infinitely many solutions.

Looking across the historical development of linear systems, one sees that the comprehension and articulation of what it means to be a solution to a linear system of equations became more refined and well articulated over time. Solution sets to linear systems were historically conceptualized first in terms of solving processes and subsequently as mathematical objects; students' understanding of solutions is likely to progress similarly. In order to develop a conceptual understanding of many of the key ideas in an introductory linear algebra course, students' similarly need to come to understand solutions to systems of linear equations as mathematical objects, particularly in the case when the solution is not unique. Over 1500 years passed between the articulation of the Chinese solution process using counting boards and the general descriptions of solution sets that were developed through the theory of determinants (and Euler's observation about inclusive dependence), so this shift from a process view of solutions to an object view of solutions is likely to be similarly non-trivial for students.

More research is needed to better understand the nature of difficulties students experience in coming to understand the solution sets to systems of linear equations, and the rationale for processes that give rise to these solution sets. However, history affords many rich insights into sources of challenge, as well as ways in which one might recreate the intellectual need for a mathematical idea in his or her classroom and pair that need with tools that position students for learning based in meaningful, historically informed mathematical activity.

ACKNOWLEDGMENTS

I extend my deepest thanks to Frank Lester, Kathy Clark, and several anonymous reviewers for help, patience, and guidance in developing this work and for the invaluable feedback provided on multiple iterations of this paper.

REFERENCES

1. Althoen, S.C. and R. McLaughlin. 1987. Gauss–Jordan reduction: A brief history. *The American Mathematical Monthly*. 94(2): 130–142.
2. Avital, S. 1995. History of mathematics can help improve instruction and learning. In F. Swetz, J. Fauvel, O. Bekken, B. Johansson, and V. Katz (Eds), *Learn from the Masters*, pp. 3–12. Washington, DC: Mathematical Association of America.
3. Boyer, C. 1985. *A History of Mathematics*. Princeton, NJ: Princeton University Press.
4. Cayley, A. 1857. A memoir on the theory of matrices. *Philosophical Transactions of the Royal Society of London*. 148: 17–37.
5. Cramer, G. 1750. *Introduction a L'analyse des Lignes Courbes Algébriques*. Geneva: Chez les frères Cramer & Cl. Philibert. Retrieved 13 May 2011, from Google Books database.
6. Dorier, J.-L. 2000. Epistemological analysis of the genesis of the theory of vector spaces. In J.-L. Dorier (Ed.), *On the Teaching of Linear Algebra*, pp. 3–81. Dordrecht, The Netherlands: Kluwer Academic Publishers.
7. Fearnley-Sander, D. 1979. Hermann Grassmann and the creation of linear algebra. *The American Mathematical Monthly*. 86(10): 809–817.
8. Freudenthal, H. 1991. *Revisiting Mathematics Education: China Lectures*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
9. Gauss, K. 1857. *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections* (C.H. Davis, Trans.). Boston, MA: Little, Brown and Company. (Original work published 1809.)
10. Gauss, K. 1957. *Work (1803–1826) on the Theory of Least Squares* (H.F. Trotter, Trans.). Technical Report No.5, Statistical Techniques Research

Group, Princeton, NJ: Princeton University. (Original work published 1811.)

11. Gravemeijer, K. 1999. How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning*. 1(2): 155–177.
12. Heeffer, A. 2010. From the second unknown to the symbolic equation. In A. Heeffer and M. Van Dyck (Eds), *Philosophical Aspects of Symbolic Reasoning in Early Modern Mathematics*, pp. 57–103. London: College Publications.
13. Katz, V. J. 1995. Historical ideas in teaching linear algebra. In F. Swetz, J. Fauvel, O. Bekken, B. Johansson, and V. Katz (Eds), *Learn from the Masters*, pp. 189–206. Washington, DC: Mathematical Association of America.
14. Kleiner, I. 2007. *A History of Abstract Algebra*. Boston, MA: Birkhauser.
15. Lesh, R. & Doerr, H. 2003. Foundations of a models and modeling perspective on mathematics teaching, learning, and problem solving. In R. Lesh and H. Doerr (Eds), *Beyond Constructivism: Models and Modeling Perspectives on Mathematics Problem Solving, Learning, and Teaching*, pp. 3–33. Mahwah, NJ: Lawrence Erlbaum Associates.
16. Mikami, Y. 1914. On the Japanese theory of determinants. *Isis*. 2(1): 9–36.
17. Philipp, R. A. 1992. The many uses of algebraic variables. *The Mathematics Teacher*. 85(7): 557–561.
18. Possani, E., M. Trigueros, J. G. Preciado, and M. D. Lozano.2010. Use of models in the teaching of linear algebra. *Linear Algebra and its Applications*. 432(8): 2125–2140.
19. Rasmussen, C. 2002. Instructional materials for a first course in differential equations. Unpublished document. Purdue Calumet University, Hammond, IN.
20. Sfard, A. 1991. On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*. 22(1): 1–36.
21. Shen, K., J. N. Crossley, and A. W.-C. Lun. 1999. *The Nine Chapters on the Mathematical Art: Companion and Commentary*. New York: Oxford University Press.
22. Swetz, F. J. 1995. Using problems from the history of mathematics in classroom instruction. In F. Swetz, J. Fauvel, O. Bekken, B. Johansson, and V. Katz (Eds), *Learn From the Masters*, pp. 25–38. Washington, DC: Mathematical Association of America.
23. Trigueros, M. and S. Jacobs.2008. On developing a rich conception of variable. *Making the Connection: Research and Teaching in Undergraduate Mathematics Education*. 73, 3–13.
24. van der Waerden, B. L. 1983. *Geometry and Algebra in Ancient Civilizations*. Berlin, Germany: Springer-Verlag.

BIOGRAPHICAL SKETCH

Christine Andrews-Larson is an Assistant Professor of Mathematics Education in the College of Education at Florida State University. Her research focuses on teacher learning, with a particular emphasis on teacher workgroups and professional networks. Her dissertation focused on student thinking, modeling, and history of mathematics with an eye toward instructional design in the context of undergraduate linear algebra. Her post-doctoral research focused on opportunities for teacher learning through site-based professional development and teachers' professional networks in the context of working to improve middle school math instruction at scale. She is currently working to coordinate these two lines of research for the purpose of understanding how to scale up inquiry-oriented instruction with a focus on post-secondary mathematics.